Tracking and imaging humans on heterogeneous infrared sensor array for tactical applications

Steven D. Feller, Evan Cull, David Kowalski, Kyle Farlow, John Burchett, Jim Adleman, Charles Lin, David J. Brady

Fitzpatrick Center for Photonics and Communication Systems and Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708

Abstract

We recently implemented a heterogeneous network of infrared motion detectors and an infrared camera for the detection, localization, tracking, and identification of human targets. The network integrates dense deployments of low cost motion sensors for target tracking with sparse deployments of image sensors for target registration. Such networks can be used in tactical applications for local and distributed perimeter and site security. Rapid deployments for crisis management may be of particular interest. This paper focuses particularly on the need for applications that deal with relatively dense and complex source fields such as crowds move through sensor spaces.

Keywords: Sensor Network, Distributed Sensing, Human Target Tracking

Introduction

Sensor networks are a key aspect of the U.S. Army's Future Combat Systems (FCS) and have been under development for over 30 years [1]. This research has led to the development of specialized components and intelligent algorithms for highly specialized systems. As the technology has improved, there has been considerable research into the theory behind sensor networks. Design topologies come in parallel, serial, and tree varieties along with ad hoc organizational structure [4]. Statistical models for determining optimal network configuration have also been developed [5,6]. In addition, there have been rapid advancements in the design of flexible, reconfigurable networks of smart sensors [2,7].

While these systems are becoming increasingly robust, the use of specialized components makes them expensive to deploy and maintain. For large sensor network deployments, this cost may be prohibitive, requiring lower cost alternatives. As sensors systems have been adapted for home and commercial security applications, simple detectors and imaging devices have become readily available. As interest in these security applications continues to increase, the diversity and availability of commercial detectors will also continue to improve, providing the opportunity for effective, low cost sensor network solutions.

In this paper we discuss a flexible sensor network developed using current off-the-shelf (COTS) components. Our focus in designing this network was to see how we could combine a large number of low-cost motion sensors with a small number of high resolution imaging components to develop a system for the detection, tracking, and identification of human targets. We implemented this system using COTS technologies to determine the effectiveness of commercially available components for sensor network applications.

For this system, we had to consider the following factors:

- *Low Power* – In order to facilitate the rapid deployment of these sensor networks in diverse environments, the system needs to be self-sufficient. For long-term applications, it is important that sensor network is operational for a considerable period of time, often in environments where external power is not available. Therefore, the power consumption of the components in these systems is a key factor in the life of the system.

- *COTS Technology* – COTS technologies are commercially produced, readily available components. Because these parts are mass-produced, COTS technologies are considerably cheaper than specially developed equipment. Commercial components are also standardized, increasing the opportunity for future extensions to existing equipment.

- *Wireless Communication* – Avoiding the overhead of providing wired connections between the components facilitates the rapid deployment of a sensor network. With the recent advances in wireless technologies, we can expect consistent, reliable, and robust communication using existing wireless devices.

- *Infrared Sensing* – These systems are designed to work under a variety of lighting conditions. By using infrared sensors over traditional light based detectors, we can ignore the lighting discrepancies. It has also been shown that infrared imaging can be sufficient for face recognition and target identification applications. [8].

- *Dense Low-Bandwidth Sensor Distribution* – One of the advantages of sensor networks lies in the relatively low cost of sensors compared to the communication and computational components of the system. By using a dense distribution of low-bandwidth sensors we can potentially reduce the computational requirements of the system and reduce the overall cost.

- *Sparse High Resolution Camera Distribution* – A small number of high-resolution cameras is required to provide the data necessary for target recognition. We reduce the number of cameras by using the low-bandwidth sensors to characterize much of the environment.

For this project, we focus on the use of COTS components. The increased diversity and availability of commercial sensors prevents this from being an excessive design limitation. Based on the capabilities of these components, we designed a wireless network of low-cost infrared motion sensors capable of locating and keeping human targets within the view of a high-resolution infrared camera. This is a low-cost, flexible system that has considerable potential in many perimeter security applications.

The initial implementation of our sensor network was deployed in a rough tobacco warehouse that recently became available to our group. Within this building we have dedicated 1500 square feet of space for the deployment of this network. The space is densely populated with columns located eight feet apart in a grid-like configuration, but is otherwise empty. Lighting is minimal at best, with only a handful of standard bulbs available within the area of interest. The space is raw with no environmental control and only a single 20-amp circuit to provide power. In addition to the physical limitations, we are also restricted from making any permanent modifications to the space. An image of the warehouse space is shown in Figure 1.

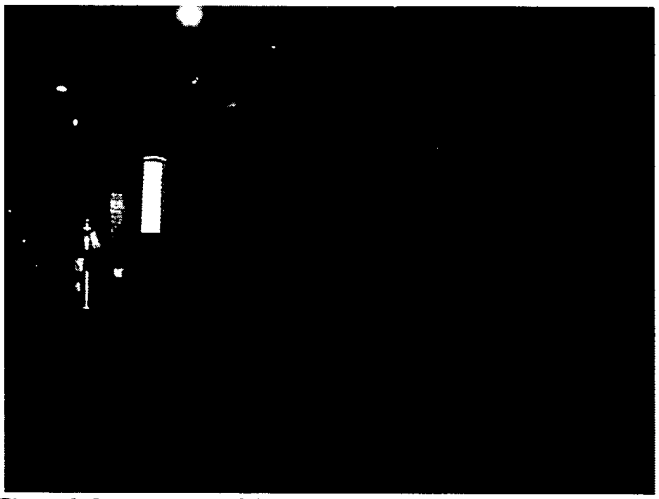

Figure 1. Interior view of the warehouse prior to network installation

This paper will discuss the design and deployment of this system and present the results from our preliminary sensor network. Section 2 describes the hardware components of the system, section 3 will discuss the localization and tracking algorithms, and section 4 describes the ability of the sensor network to track human targets. Section 5 describes future plans for this research.

Sensor Network Design

The sensor network we discuss in this paper was developed to be a low cost, low power, easily deployable system capable of accurately tracking human targets. Although we focused on developing a flexible network, there were several environmental constraints that had to be taken into consideration. In this section, we discuss the design of the system.

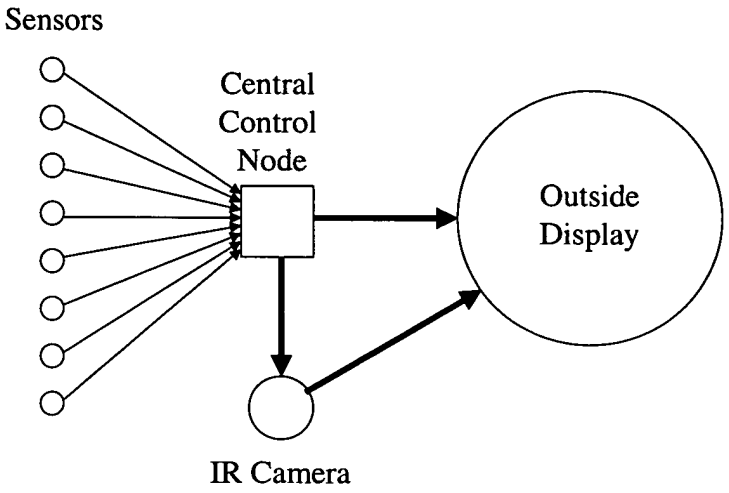

Figure 2. Logical sensor network diagram

The basic concept behind sensor networks is the use a large number of simple, relatively low cost sensors to provide a better analysis of the environment than a single sensor can measure. For the detection, localization, and tracking of human targets, we use twenty-two motion detectors spread across the imaging space with known location and orientation. A central control node receives sensor data and fuses the data to extrapolate the location of events in the environment. We use the location information computed on the control node to keep a high-resolution infrared camera pointed towards the target for identification purposes. The location data from the control node and the data from the IR camera are then streamed to a remote location where it is displayed. Figure 2 shows a logical diagram of our sensor network.

The motion sensors used in this system are binary infrared detectors commercially available over the internet at minimal cost. These sensors transmit data to the control node using the X10 communication protocol. This particular model can support up to 256 uniquely identified devices. Once deployed, a sensor remains in standby, waiting to detect motion. The sensor consists of two adjacent pyroelectric diodes. If one diode is activated, and the adjacent diode is subsequently tripped within a certain time interval, the sensor reports a motion event. The sensor has a refractory period of approximately one second after reporting an event before it will report another event to prevent multiple indications of the same signal. The duration of the refractory period is adjustable on each individual sensor. In areas of decreased sensor density or increased field of view, it is desirable to decrease the refractory period as much as possible. When motion is detected, the sensor transmits its unique ID to the controlling computer to indicate the detection. This signal on detection paradigm with minimal processing allows these sensors to run several months consecutively on two AA batteries. An image of a motion detector can be seen in Figure 3.

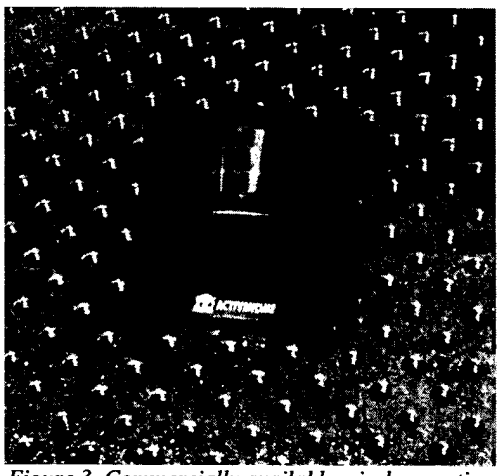

Figure 3. Commercially available wireless motion

The data from each of the motion detectors in the network is received at the central control node. This device uses the asynchronously received data from the triggered sensors and previously determined sensor orientation and location to extrapolate the location of a target. With our emphasis on low power consumption and rapid deployment, we selected a PC104 modular computer with a Pentium processor for the control node. The PC104 was chosen for its flexibility and small size, as can be seen in Figure 4. Ideally, we would like to find or develop a central control system that is able to operate in a low power consumption state while waiting for an event, yet still be able to wake up and perform necessary computations when an event occurs.

Figure 4. PC104 Control Node

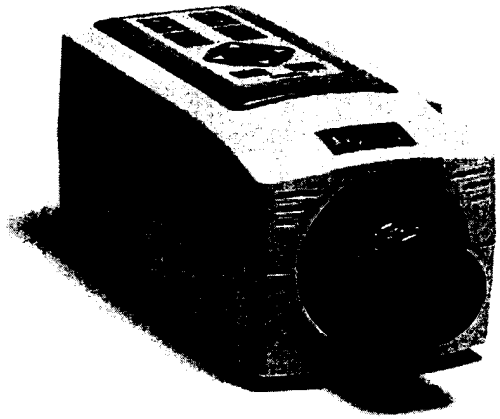

Figure 5. Infrared Camera

For high-resolution imaging, we use a Merlin near infrared camera offered by Indigo Corp. This scientific camera produces a 320x256 NTSC video signal. We mounted the camera on an automated rotation stage, which was controlled through a computer. This computer receives target location data from the control node and converts the coordinates into the corresponding angle on the rotation stage. Figure 5 is an image of the infrared camera.

Since our research interests lie in providing the data necessary for target identification and not in the identification process itself, the infrared camera and its control system are a passive observer in the sensor network and do not participate in target localization. The camera controller only receives coordinates of the detected target from the control node and automatically points to the specified location. As the target moves through the space, the control node continuously redirects the camera to the new location of the target. We use our ability to keep the target in view of the camera as a metric for determining success.

Communication between the motion detectors and the central control node occur in the unregulated 2.4 GHz band. Each detector has a built in transmitter while the controlling PC104 receives these signals through an antenna connected through its serial port. The control node and camera controller interact over the 80211.b networking protocol using a standard wireless access point. The total bandwidth of the system is minimal, with each detector sending only a few bytes per detection and communication between the camera controller and the central node reduces to simple coordinates.

Algorithms

This sensor network is designed to work at several levels of complexity. The simplest functionality involves detecting motion within the environment. This can be easily achieved within minimal resource expenditure, and is specifically what motion detectors are designed to do. As we attempt to determine the source of the motion and track a target the difficulty of the problem and the resources required to achieve these goals are greatly increased, but remain tractable. The problem of target identification is a difficult challenge that is outside of the scope of our research. Instead of performing the identification process, we focus on extracting the data necessary to perform this kind of analysis.

Once motion has been detected, we begin the process of determining the location of the source. If the distribution of sensors is sufficiently dense, any motion will trigger multiple detections. With prior knowledge of sensor location, orientation, and field of view we are able to determine the regions with the highest probability of containing a target using a two-dimensional backprojection algorithm. This algorithm creates a two-dimensional pixel map that is spatially consistent with the physical environment. When a motion sensor is triggered, the pixel values corresponding to the field of view of the detector are incremented. The intensity of pixels where the zone of detection for two activated sensors overlap will be greater than the pixels that are affected by fewer active sensors. The highest intensity pixel values in the

pixel map represent the areas with the greatest probability of finding a target. The coordinates of the highest intensity points are used to determine the correct angle of the camera.

This backprojection technique was chosen because of its flexibility and extensibility. If a sensor is added to the network or an existing sensor is relocated, we only need to add the orientation and location parameters of the sensor to contribute to the system. Since the field of view can be uniquely specified for each sensor, we are able to incorporate a wide range of diverse sensors into the system with minimal effort. The backprojection algorithm also allows us to weight the importance of each sensor in the network. By weighting the value of the detectors, we can produce a network that focused on detections in certain areas over others.

As a target moves through the space, the area of interest in the pixel map needs to reflect this change. When a target leaves a region, we need to reduce the importance of that space while increasing the importance of adjacent spaces. This is accomplished by constantly reducing the intensity of the pixels in the pixel map over time. This fading approach permits a sensor that detects a single event to only affect the space for a short duration. We can track a target by recording the location of the highest intensity value over time.

We designed two methods to allow users to view where events occur within the sensing environment. For data received directly from the sensors, we developed an OpenGL graphical user interface that shows a three-dimensional representation of the environment. Within this model, a cone projecting from the sensor location represents the field of view for each sensor. When a detector is activated, the cone representing its field of view is highlighted. By examining the intersection of multiple active cones a user is able to determine the location of the target within the space. The second method for users to view the location of targets simply involves displaying the video stream generated by the infrared camera. This provides a qualitative indication of where targets are in the space.

Results

The installation of the sensor network was completed in a matter of hours. We placed the motion detectors on the existing columns within the space. This was a choice of convenience, since the algorithms we use are designed to handle random sensor distributions. The central control node was placed near one of the power outlets coincidentally located near the center of the system. The physical location of the control unit is not particularly important as long as it is within the transmission range of each of the detectors. The camera for tracking targets was arbitrarily placed along the edge of the network with a fairly clear view of the space. Figure 5 shows a layout of the sensor space.

When a target enters the sensor network it is almost instantly detected and the infrared camera automatically turns in the general direction of the event. The camera is only able to roughly determine where an object is located until multiple sensors are triggered. Once two detectors with intersecting field of view are triggered, the sensor network is able to locate the target object within the range of view of the camera. The network is able to recognize changes in the environment well enough for the camera to continue tracking a target while it is in the sensing environment. Figure 7 is an example of the graphical user interface when a target is located in the top left corner of the environment. Figure 8 is an infrared image of a target walking through the environment.

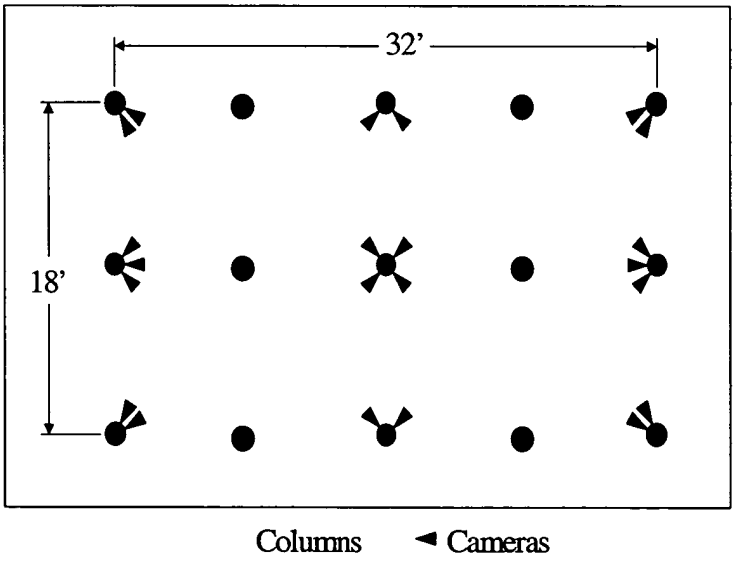

Figure 6. Layout of the sensor space.

Figure 7. Graphical interface with a target located in the top left of the space.

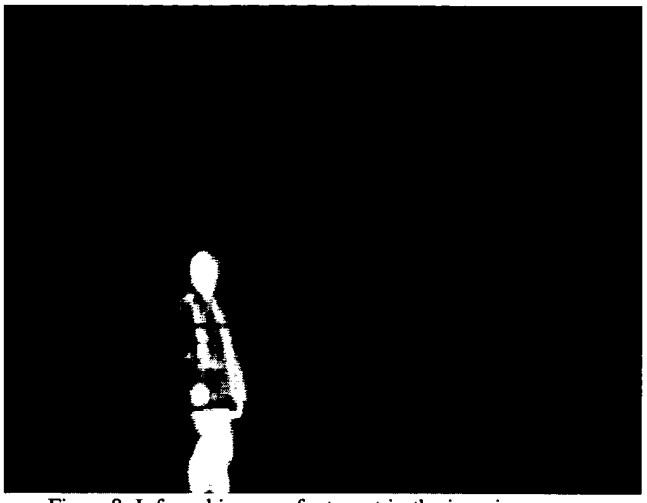

Figure8. Infrared image of a target in the imaging space.

The target tracking for the network works well. When multiple sensor detect a target, the infrared camera is able to get at least most of the target in its field of view, although the image is not necessarily centered. When multiple sensor detect a target, the infrared camera is able to get at least most of the target in its field of view, although the image is not necessarily centered. The graphical display shows which sensors have been triggered but does not provide any further processing. This user is easily able to extrapolate the location of a target from the intersection of cones displayed in the user interface.

If more than one target enters the sensor space, the control node is able to track their location, but not necessarily the number of targets. This is where the sensing capability of the motion detectors becomes a limiting factor. They are unable to determine if a single person is making motion or if multiple targets in close proximity are triggering events. While the infrared camera may be able to provide some help in this determination, if more than one group of targets is in the space, we will only be able to focus on a single group. We can incorporate additional cameras into the sensor network to watch multiple groups, but the number of disparate groups we can observe will still depend on the number of cameras.

The communication between the components of the networks worked seamlessly with only basic installation. The motion detectors communicate with the control center using their own proprietary protocol and communication between the control center and the camera controller is achieved over a standard 80211.b access point. The limited amount of data being transmitted across the entire system indicates that a much denser distribution of detectors can be supported.

This system was designed using COTS technologies that are commercially available. Using these components, we were able to achieve our goal of accurately detecting, localizing, and tracking targets and to produce the data necessary to identify a target. The sensing components of the network use infrared detection to eliminate the lighting issues that surface in traditional imaging problems and take advantage of the heat generated by human targets. By choosing components with wireless communications and low power consumption, we were able to develop this network for rapid deployment. With this combination of components, we achieved our goal of developing a low-cost, rapidly deployable, effective system for analyzing human targets.

Future Work

The sensor network presented in this paper was developed to determine how existing technologies can be applied to the detection, localization, tracking, and identification of human targets. There is a considerable amount of research that remains to be done to explore the capabilities and potential of this base system.

In order to understand the capabilities of this system we need to develop techniques to determine the effectiveness of the array with better accuracy. We need to perform formalized tests to determine the response of the network to a variety of targets in a wide range of environments. This will allow us to determine the limits on how many targets can be tracked in the system and how the accuracy degrades as the number of targets increases. We also need to test the networks ability to differentiate between human and non-human targets.

A statistical analysis of the system will allow us to predict overall coverage and determine the resolution of the system. The motion detectors used in the network discussed in this paper all have the same range of view. By changing the angular range of these sensors we may be able to provide greater coverage with equal or potentially even better accuracy than the current system. An analysis of the reliability of the sensors will allow us to predict component failure and determine the amount of redundancy that is necessary in the system. This statistical model will allow us to improve the accuracy and robustness of the system.

The use of self-registering sensors can significantly reduce the time required to deploy the network. For this implementation, we had to accurately place and align each of the motion detectors in the sytem, which required a considerable amount of time. By designing sensors that can determine their location and orientation automatically, we can significantly reduce the time of deployment. Eventually, we would like to be able to simply throw a collection of these sensors into the environment, have each sensor identify its own coordinates, and begin tracking targets. By examining network coverage, we can manually place a few sensors in low sensor density locations.

The tracking algorithms can be extended to provide three-dimensional tracking. While this is not necessary for many applications, there are environments where the targets are not necessarily located at the same level. Three-dimensional tracking will provide a better estimation of target location and can be used to better direct the high-resolution cameras.

Future network will use a hierarchical detection protocol for managing power consumption across the network. In this kind of system, only a few detectors will be active while there are no registered events in the environment. The rest of detectors and nodes on the system will reside in a dormant, low-power mode. If one of the global sensors detects an event, it will wake up other sensors in the network as well as the central control. This second tier of sensors will provide a higher resolution measurement of the environment, which the control center uses to extract the location of the target. When the location of a target is well defined, the high-resolution camera will be activated to image the environment. This approach will greatly reduce the power required by the system and reduce unnecessary communication between the components.

A sensor network needs to be capable of abstracting the massive amounts of data measured by a dense heterogeneous mix of sensors into a manageable stream of information. The components of such a system will require a balance between computational capabilities, communication bandwidth, and power consumption. Based on these concepts, we developed a low-cost sensor network for the detection, localization, and tracking of human targets. This network will provide a basis for further research into the capabilities of sensor systems. While there is still much work to be done, the system we developed has demonstrated the potential of dense deployments of low-cost motion sensors coupled with sparse deployments of high-resolution imaging components.

References

1. J. Nemeroff, L. Garcia, D. Hampel, S. DiPierro, "Application of sensor network communications", Military Communications Conference, 2001, Communications for Network-Centric Operations: Creating the Information Force, IEEE. Volume: 1, 2001, pp. 336-341.

2. M.G. Corr and C. M. Okino, "Networking Reconfigurable Smart Sensors", Proceedings of SPIE: Enabling Technologies for Law Enforcement and Security, vol. 4232, Nov. 2000.

3. B. Jones, "Design of a remotely operated intrusion detection system for security applications", 1993 International Carnahan Conference on Security Technology, Proc. IEEE, pp. 145-153.

4. R. Viswanathan, P. K. Varshney, "Distributed Detection with Multiple Sensors: Part I – Fundamentals", Proceedings of the IEEE, Vol. 85, No. 1, January 1997.

5. D. E. Brown, J. B. Schamburg, "A Simulation-Optimization Methodology for Sensor Placement", Systems, Man, and Cybernetics, 1997. IEEE Conference on Computational Cybernetics and Simulation 1997. Vol. 1, pp. 439-443.

6. S. Meguerdichian, F. Koushanfar, M. Potkonjak, M. B. Srivastava, "Coverage problems in wireless ad-hoc sensor networks", INFOCOM 2001. Proceedings IEEE, Vol. 3, pp. 1380-1387.

7. J. M. Kahn, R. H. Katz, K.S. J. Pister, "Mobile Networking for Smart Dust", ACM/IEEE Intl. Conf. on Mobile Computing and Networking (MobiCom 99), Seattle, WA, August 17-19, 1999.

8. F. Prokoski, "History, current status, and future of infrared identification", Computer Vision Beyond the Visible Spectrum: Methods and Applications, 2000. Proceedings. IEEE Workshop on, 2000, pp. 5-14.